国外深度伪造技术政策文本分析与启示
——以美国、欧盟为例

赵雪芹，李天娥，胡慧慧

（湖北大学 历史文化学院，武汉 430062）

摘　要：［目的 / 意义］深度伪造技术政策是技术异化治理的重要手段，通过分析美国与欧盟深度伪造政策内容，从而思考中国如何运用政策手段来对深度伪造技术异化进行治理。［方法 / 过程］文章选择美国与欧盟发布的 21 份政策文献作为研究对象，利用 Nvivo12 软件进行政策编码。然后以基本政策工具作为 X 维度，以技术异化治理维度作为 Y 维度，综合 X&Y 维度进行内容分析。［结果 / 结论］研究发现，美国与欧盟的深度伪造技术政策以环境型政策为主且较为关注对技术、个体和组织的治理，除此之外，也关注技术伦理的宣传与教育。据此，得出中国政府应制定深度伪造技术异化治理方案、促进政府主导的深度伪造技术创新实践、实现技术伦理与政策工具的结合、积极参与国际深度伪造技术异化治理等启示，从而建立健全中国的技术异化治理政策体系。

关键词：深度伪造；政策分析；技术异化；国家安全；虚假信息

中图分类号：D521；G255　　**文献标识码：**A　　**文章编号：**1002-1248（2022）09-0060-12

引用本文：赵雪芹，李天娥，胡慧慧．国外深度伪造技术政策文本分析与启示——以美国、欧盟为例［J］．农业图书情报学报，2022，34（9）：60-71.

1 引　言

近年来，随着数字技术不断更新迭代，人工智能技术应用不断推陈出新，尤其是以智能化算法为基础的深度学习使得人们的生活发生了巨大改变，其中较具代表性的便是 Deepfake 技术。“Deepfake”由“Deep Learning”和“Fake”结合而成，即“深度伪造”，也称“深度造假”，是指运用深度学习算法去创建或修改音视频、图像等内容，最终伪造一些虚假的信息和内容[1]。深度伪造技术是人工智能技术发展过程中自然而然出现的衍生技术，其核心是深度学习算法中的生成对抗网络（GAN）算法，该算法可以对语音、图像、文字等信息内容进行修改。目前在日常生活中，语音合成、AI 换脸等深度伪造技术应用已十分常见，FaceSwap，Zao，Lyrebird 等移动 APP 使用户可以自由地创建并发布内容。然而，技术一直以来都具有双刃剑性质，一方面可以创造价值，造福人类；另一方面具有潜在的安全隐患，危害个人、社会、国家安全，深度伪造技术也不例外。

收稿日期：2022-07-05

基金项目：国家社科基金一般项目“移动图书馆用户信息交互行为中的情感体验研究”（18BTQ061）

作者简介：赵雪芹（1983-），女，博士后，教授，研究方向为信息服务。李天娥（1997-），硕士研究生，研究方向为档案信息组织。胡慧慧（1998-），硕士研究生，研究方向为档案信息服务

目前，国内外关于深度伪造的研究主要集中在以下主题：一是针对深度伪造的检测和溯源技术研究，如基于循环神经网络的深度伪造视频检测[2]、基于图像纹理特征的深度伪造图像检测算法[3]、深度伪造图像追踪溯源方法[4]等；二是深度伪造所带来的困境研究，如使政治舆论复杂化[5]、威胁国家安全[6]、破坏社会稳定及民主[7]等；三是深度伪造治理研究，对深度伪造技术应用过程中的潜在刑事治理风险[8]、治理模式[9]进行探究，并从技术角度提出规制深度伪造的方案[10]。由此可知，为引导深度伪造技术向善发展，国家必须对深度伪造技术主体的行为进行有效管理和限制，不能任其自由发展。然而，中国目前尚未有针对深度伪造技术的政策，因此，本文采用文本分析法，对美国和欧盟的深度伪造政策文本进行分析，得出他们治理深度伪造技术的一般手段，进而为中国深度伪造技术的治理提出政策建议，以期为中国深度伪造技术以及其他人工智能技术的发展提供有力的政策保障。

2 深度伪造技术的威胁

当前，深度学习算法推动了数字经济的发展，成为新的行业增长点。但同时，由深度学习算法而衍生的深度伪造技术，对网络安全造成严重的威胁，加剧网络安全风险。2020 年，澳大利亚与美国分别发布《深度造假武器化——国家安全与民主》报告[11]与《Deepfakes：基本威胁评估》报告[12]。根据这两个报告，可以总结深度伪造技术带来的威胁主要在 3 个方面：一是在国家层面，深度伪造技术降低了国家之间信息站的成本，恶意行为主体会利用深度伪造技术制造虚假信息来引发舆论，进而给国家安全带来不良影响，如 2016 年有人发布深度伪造内容干预了美国总统大选；二是在社会层面，深度伪造技术逐渐成为网络犯罪者的工具，如在 2019 年 3 月，英国有犯罪分子利用深度伪造技术成功诈骗 22 万欧元；三是在公众层面，深度伪造技术的广泛使用可能会破坏人们与个体或政府机构之间的正当沟通机制，大量虚假的合成内容会削弱公众对于信息真伪的辨认力，从而降低人们对政府制度的信任度，引发信任危机。如 2019 年 6 月，美国皮尤研究中心发布报告，称超过 1/3 的调查者认为“虚假新闻”导致他们不能接收真正的新闻[13]。

综上所述，可知深度伪造技术的滥用对国家安全、社会公共安全、个人隐私和声誉均造成了不同程度的威胁，因而有必要从政策层面对深度伪造技术的开发、使用等行为进行约束。首先，深度伪造技术政策可帮助政府更好治理网络环境，使执法机关能够有法可依，进而更加精准地打击网络犯罪；其次，深度伪造技术政策能够引导技术主体修正自身行为，不再继续借助技术损害他人权益；最后，对于技术发展而言，政策的引导和限制是极其必要的。

3 深度伪造政策文本选择及分析框架构建

美国与欧盟已关注到深度伪造技术的异化现象，且制定了一系列的政策加以治理，对中国深度伪造技术的治理具有一定的借鉴意义，因此笔者选择美国与欧盟的深度伪造技术政策文本进行分析。

3.1 政策文本选择

为保证研究的代表性与权威性，政策文本的选择将根据以下原则所进行：第一，政策发文机构必须为美国的联邦政府、州政府或欧盟委员会；第二，仅收集已在政府网站中公开发布的政策；第三，所收集的政策主题需与深度伪造紧密相关，因此本文以“Deep Fake”“Deceptive”“Disinformation”“Deep Fakes”等作为检索关键词，在美国的国会网站、各州政府网站及欧盟委员会网站进行政策收集。基于以上原则，收集时间为 2022 年 1 月 20 日—2022 年 1 月 31 日，初步获得美国及欧盟的共 25 份相关政策文本。下一步，需要对政策文本进行整理，首先深入了解每一份政策文本的主题，剔除与深度伪造技术无关的文本；其次，对所有政策文本进行扫描和识读，把与本研究相关的政策内容汇总成一个文档，比如，在《2020 年国防授权法案》中，只需抽取其中与深度伪造相关的两个章

节；最后，经过多次检查与核对，整理得到 21 份政策文本（表 1），以及一份 66 837 字符的文档，以此作为本文的研究材料。

3.2 政策分析框架

本文采取二维分析框架对深度伪造政策文本进行分析，首先以基本政策工具作为 X 维度，然后以技术异化治理的角度作为 Y 维度，最后将二者结合进行综合分析，即研究政策工具在技术异化治理上的应用情况，分析框架如图 1 所示。

3.2.1　X 维度：基本政策工具

政策工具是政府达成某种目标的一种重要手段，在政府制定政策目标时，政策工具的科学设计与组合是实现政策目标的前提[14]。因此，可通过分析政策中的政策工具类型及其组合运用去研究政策的合理性与科学性。目前，政策工具在学界有许多分类，如

表 1　美国与欧盟深度伪造政策文本

Table 1　Deepfake policy text in the US and EU

发文机构	时间	政策文本
美国国会	2019 年 7 月	《2019 年“深度伪造”报告法案》（The Deepfake Report Act of 2019）
	2019 年 7 月	《2020 年国防授权法案》（The National Defense Authorization Act for Fiscal Year 2020）
	2018 年 12 月	《2018 年恶意伪造禁令法案》（Malicious Deep Fake Prohibition Act of 2018）
	2019 年 3 月	《2019 年商业人脸识别隐私法案》（CFRPA2019）
	2019 年 4 月	《隐私权利法案》（Privacy Bill of Rights）
	2019 年 6 月	《深度伪造责任法案》（Deepfakes Accountability Act）
	2019 年 9 月	《生成对抗性网络法案》（Identifying Outputs of Generative Adversarial Networks (IOGAN) Act）
	2019 年 6 月	《2018、2019 和 2020 财年达蒙·保罗·纳尔逊和马修·杨·波拉德情报授权法案》(Damon Paul Nelson and Matthew Young Pollard Intelligence Authorization Act for Fiscal Years 2018, 2019, and 2020）
	2019 年 5 月	《要求国防部长对武装部队成员及其家属的网络剥削和其他目的进行研究的法案》（A Bill to Require the Secretary of Defense to Conduct a Study on Cyber exploitation of Members of the Armed Forces and Their Families, and For Other Purposes）
马萨诸塞州政府	2019 年 1 月	《防止深度造假被用来促进犯罪或酷刑行为的法案》（An Act to Protect Against Deep Fakes Used to Facilitate Criminal or Torturous Conduct）
纽约政府	2018 年	《公民权利法修正案》（An Act to Amend the Civil Rights Law）
德克萨斯州政府	2019 年 6 月	《关于制作欺骗性视频意图影响选举结果的刑事犯罪法案》（Relating to the Creation of a Criminal Offense for Fabricating a Deceptive Video with Intent to Influence the Outcome of an Election）
弗吉尼亚州政府	2019 年 7 月	《非法传播或出售他人影像》（Unlawful Dissemination or Sale of Images of Another Person）
加利福尼亚州政府	2019 年 2 月	《犯罪：欺骗性记录》（Crimes: Deceptive Recordings）
	2019 年 10 月	《选举：欺骗性的音频和视觉媒体》（Elections: Deceptive Audio or Visual Media）
	2019 年 10 月	《使用数字或电子技术描述个人:色情材料:诉因(加利福尼亚州 AB-602)》（Depiction of Individual Using Digital or Electronic Technology: Sexually Explicit Material: Cause of Action（Calif. AB-602））
欧盟委员会	2018 年 4 月	《应对线上虚假信息：欧洲方案》（Tackling Online Disinformation: A European Approach）
	2018 年 5 月	《通用数据保护条例》（General Data Protection Regulation）
	2019 年 5 月	《非个人数据自由流动条例》（Free Flow of Non-Personal Data Regulation）
	2019 年 6 月	《网络安全法案》(EU Cybersecurity Act）
	2018 年 9 月	《反虚假信息行为准则》（Code of Practice on Disinformation）

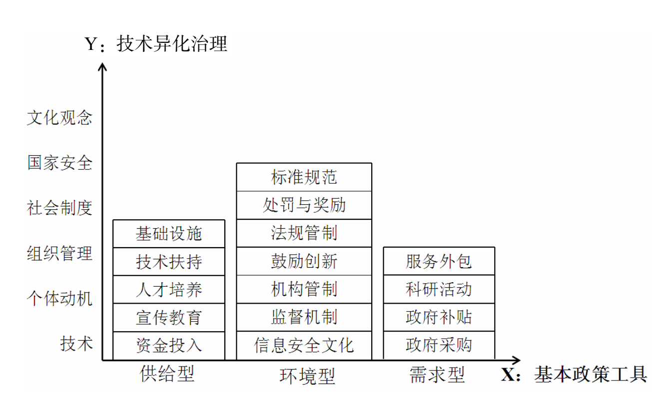


图 1　深度伪造技术政策分析框架

Fig.1　An analytical framework for deepfake technology policy

HOWLETT 等将政策工具分为信息型、经济型、权威型和志愿型[15]；PHHAL 等将政策工具划分为自愿型、强制型、混合型[16]。虽然政策工具分类方式多样，但在进行政策文本分析时，研究者可依据具体研究对象和研究问题进行合理的选择。本文借鉴 ROTHWELL 和 ZEGVELD 的划分，并以此作为政策分析的 X 维度，从政策对技术影响的角度将政策工具划分为供给型、环境型和需求型[17]。

供给型政策工具是指政府通过提供基础设施、技术扶持、人才培养、宣传教育、资金投入等手段来保障深度伪造技术的可持续发展，此阶段最重要的是要培养技术创新者的科技向善理念及人文主义观念，在技术开发之前就有效规避其不良影响。环境型政策工具是指政府为技术创新创造有利条件，如加大对深度伪造内容的监督力度、建立技术使用反馈机制，本文所涉及的政策工具包括标准规范、处罚与奖励、法规管制、鼓励创新、机构管制、监督机制、信息安全文化。需求型政策工具主要是用于降低某种技术的市场风险，从而促进该技术的可持续发展，其重点是要给予技术创新一定的关注与扶持，尽量降低不良因素对技术创新的影响，防止技术异化。它具体包括服务外包、科研活动、政府补贴、政府采购等工具。

3.2.2　Y 维度：技术异化治理

技术异化是当前新兴科技发展演进中出现的一种现象，即在技术至上和功利主义的主导下，违背“以人为本”的科技发展初衷，违背法律、伦理和人文精神，以获取非正当收益为主要目标的一种不利于国家、社会、个人可持续发展的现象[18]。当前，科学技术无疑是国家发展和社会进步的重要基石，因而技术创新应将“以人为本”作为出发点，以社会的可持续发展作为根本目的。然而，由于目前新兴技术层出不穷且使用门槛较低，导致技术异化现象屡屡发生。深度伪造技术便是当前人工智能技术异化的典型代表，若不采取相应措施去规避深度伪造技术存在的潜在风险，任其自由地发展下去将会对国家、社会和个人造成不可逆转的影响。

为了对美国与欧盟的深度伪造政策文本进行深入的分析，本文将技术异化治理作为政策分析的 Y 维度，以探究深度伪造技术治理的情况。根据技术异化的原因，可将技术异化治理从微观到宏观分为 6 个角度，即技术、个体动机、组织管理、社会制度、国家安全和文化观念[19]，如图 2 所示。

具体而言，技术异化的原因，在技术层面体现为技术的应用场景和组合方式不相适应，以及不断增长

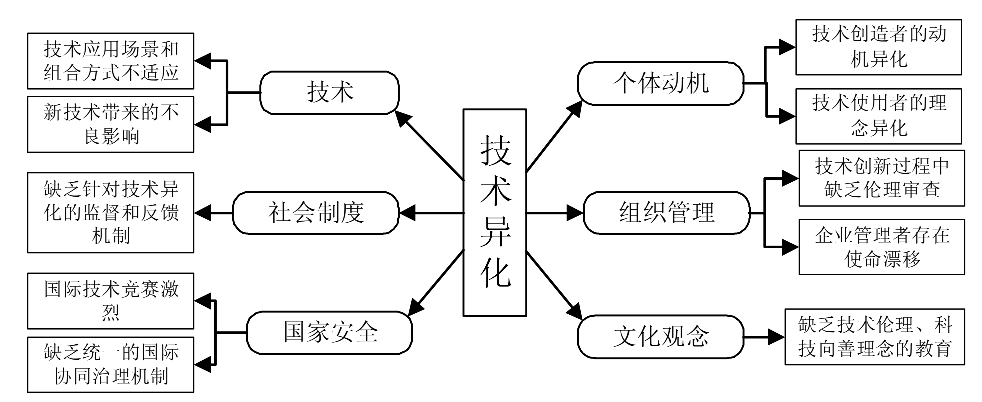


图 2　技术异化治理的 6 个角度

Fig.2　Six perspectives of dealing with technology alienation

的新技术对社会的反噬，即新技术带来的不良影响。在个体动机上体现为技术创造者的动机异化，即坚持技术至上和功利主义价值观[20]；以及技术使用者的理念异化，即没有正确的技术防控理念。在组织管理的层面表现为技术创新过程中缺乏伦理审查机制，即未考虑新技术可能造成的伦理问题；以及企业管理者存在“使命漂移”现象[21]，即没有形成规范的、闭合的管理体系。在社会制度层面表现为缺乏针对技术异化的监督和反馈机制，使得技术异化的现象愈加严重。在国家安全层面，由于技术创新可以提升国家的整体竞争优势，因此技术竞赛一直存在于国家之间[22]，并且将愈演愈烈，但是目前没有统一的国际协同治理机制，这将会影响人类命运共同体的构建，最终可能会导致技术异化在全球范围内发生，进而造成不可估计的负面影响。最后，从更加宏观的层面来看，技术异化的产生与技术文化观念的教育缺失密不可分，主要原因是在技术创新过程中，社会和国家缺乏技术伦理、科技向善、人文主义等理念的教育。综上所述，技术异化产生的原因错综复杂，所以其治理也需要从原因的角度展开，而政策制定是国家重要的治理手段之一，因此本文将技术异化治理角度与政策工具结合进行文本分析，以探究深度伪造政策是否合理、规范。

4 国外深度伪造政策文本分析

本文采取 Nvivo12 软件作为政策分析工具，首先对政策文本进行逐一阅读，了解大致内容；其次，利用 Nvivo12 软件完成政策工具的编码，根据选定的政策工具类型对政策条款逐一进行编码；最后，对编码结果进行分析，并得出结论。

4.1 政策主题分析

在对内容进行分析之前，笔者利用 Nvivo12Plus 软件分别对美国及欧盟的深度伪造政策主题进行了统计分析，经过停用词、词长度、显示字词等词频条件的设置，最终得到如图 3 所示的词语云。字体越大表示该词在政策中出现的频次越高，由图 3 看出，美国与欧盟的深度伪造政策侧重点有所不同，美国最为突出的 3 个词是“individual”“person”“act”，表明其较为关注技术对个人信息所造成的威胁，严格治理危害个人安全的深度伪造内容，从源头上杜绝深度伪造技术对个人隐私的侵犯。欧盟最为突出的 3 个词是“ict”“data”“cybersecurity”，说明欧盟通过对通信技术、数据安全、网络安全的治理，进而间接治理深度伪造



图 3 深度伪造政策词语云

Fig.3 Word cloud for deepfake policies

技术，原因是深度伪造内容依赖于各种通信技术进行传播，并对数据安全与网络安全造成影响。

4.2 X 维度分析

采用 Nvivo12 软件中的编码功能对 21 份政策文本进行编码，首先由两位成员进行逐一编码，然后再由第三人对两份编码结果进行逐一校对，进行两轮编码之后最终得到 236 个编码点，编码示例如表 2 所示。这些编码点在政策工具上的分布如表 3 所示。

由表 3 可知，供给型政策工具共有 36 个参考点，文字覆盖率为 2.47%。其中，技术扶持和宣传教育是使用较多的政策工具，技术扶持体现为政府明文指出要研究和开发可以对抗并打击深度伪造的技术，而且政府要为研究此类技术的机构提供支持。比如美国加利福利亚州在《犯罪：欺骗性记录》法案中提出资助加州大学 2 500 万美元用于研究新技术，以识别和打击 Deepfake 技术的不当使用。而宣传教育政策工具主要体现为政府为培养公众辨别数字内容真实性的能力所制定的相关措施，在很大程度上，公民数字能力的培养可以从技术使用者层面上减少深度伪造技术的异化现象。因而，宣传教育也是促进深度伪造技术持续性发展的重要政策工具之一。

环境型政策工具共有 177 个参考点，文字覆盖率为 20.31%，是使用最多的政策工具类型。其中，法规

表 2 深度伪造政策编码示例

Table 2 Coding example of deepfake policy

政策条文	政策工具	编码
……促进公共和私营部门之间以及私营部门内部的合作，特别是支持保护关键基础设施	基础设施（BI）	BI-1
……酌情研究和开发检测或以其他方式对抗和打击深度伪造和其他高级图像处理方法的技术……	技术扶持（TS）	TS-1
……ENISA应进一步加强其技术和人力的能力和技能	人才培养（TD）	TD-1
……研究教育公众辨别数字内容真实性的最佳实践	宣传教育（PE）	PE-1
从普通基金拨出2 500万美元给加州大学……	资金投入（FS）	FS-1

表 3 基于 Nvivo12 的政策工具编码统计表

Table 3 Statistical table of policy tool coding based on Nvivo12

工具类型	政策工具	参考点/个	参考点总计/个	覆盖率/%	覆盖率总计/%
供给型	基础设施	5	36	0.25	2.47
	技术扶持	11		0.83	
	人才培养	4		0.27	
	宣传教育	10		0.75	
	资金投入	6		0.37	
环境型	标准规范	35	177	4.95	20.31
	处罚与奖励	19		3.11	
	法规管制	49		5.93	
	鼓励创新	13		1.03	
	机构管制	28		1.82	
	监督机制	29		3.10	
	信息安全文化	4		0.37	
需求型	服务外包	4	23	0.33	1.89
	科研活动	13		1.02	
	政府补贴	2		0.27	
	政府采购	4		0.27	

管制工具使用最多，主要原因是美国与欧盟对深度伪造制定的政策多为法案，因此其中的法规管制条款占比较大。比如，美国的弗吉尼亚州颁布的《非法传播或出售他人影像》是针对深度伪造刑事犯罪的刑法，内容明确且具有强制性。除法规管制之外，标准规范、机构管制与监督机制 3 个是使用较多的政策工具，标准规范工具用于制定与深度伪造技术相关的定义、标准，以明确生活中的哪些行为属于深度伪造犯罪，帮助执法机关落实法律法规。机构管制工具主要是用于对相关机构的合作、技术开发、信息共享等活动进行管理和控制，以形成一个良好的技术创新环境。监督机制是政府治理深度伪造技术异化的关键手段，主要是因为深度伪造技术是否向善往往取决于技术开发者和使用者的主观行为，不能一味地采取强制性手段去打击深度伪造技术本身，因而需要监督和审查技术的开发与使用者，以便能够及时发现技术异化现象。此外，政府还采用了处罚与奖励工具，以打击深度伪造犯罪行为，并激励个体和组织的技术创新行为。

需求型政策工具共有 23 个参考点，文本覆盖率为 1.89%，是最少使用的工具类型。而在 23 个参考点中，科研活动工具占据一半以上，说明目前美国与欧盟政府都致力于研究新技术来对抗深度伪造技术，比如开发检测视频、音频文件或照片的工具，助力司法实践中的数字取证，以识别和分析数字证据的真伪。

4.3 X & Y 维度分析

在 X 维度的统计分析之上，结合 Y 维度，统计政策工具中分别涉及了技术异化治理的哪个角度，统计结果如表 4 所示。由表 4 可知，X 维度在 Y 维度各层面上的运用差距较大。在技术层面，政府使用最多的政策工具是标准规范，主要是用于界定深度伪造技术的定义、内容、使用规范等，或是界定哪些行为属于深度伪造犯罪，有助于执法部门发现深度伪造技术异化现象。

在个体动机层面，使用最多的政策工具是法规管制，政府一方面通过法规管控技术创新者的恶意行为；

表 4　X&Y 维度统计表

Table 4　Statistical table of X&Y dimension

X维度 \ Y维度	技术	个体动机	组织管理	社会制度	国家安全	文化观念
基础设施	0	2	2	1	0	0
技术扶持	4	1	2	0	3	0
人才培养	0	2	0	1	1	0
宣传教育	0	3	0	0	0	7
资金投入	2	3	1	0	0	0
标准规范	20	12	3	0	0	0
处罚与奖励	0	9	1	8	0	0
法规管制	8	22	9	3	4	0
鼓励创新	1	5	2	2	3	0
机构管制	1	1	17	1	6	2
监督机制	2	0	4	17	6	0
信息安全文化	0	0	0	0	0	4
服务外包	0	3	1	0	0	0
科研活动	0	13	0	0	0	0
政府补贴	0	0	0	0	2	0
政府采购	0	0	3	1	0	0
小计	38	76	45	34	25	13

另一方面以条例保护技术使用者的隐私权、诉讼权等公民权益不受侵害。此外，科研活动工具全部作用于个人层面，说明对于深度伪造技术异化的治理，重点需要对技术创新者的科研行为进行积极引导。在组织管理层面，政府运用最多的政策工具是机构管制，通过对深度伪造技术相关机构进行依法管理，以促进机构间的相互协同合作，实现对深度伪造技术异化的多元协同治理。在社会制度层面，使用最多的政策工具是监督机制，该工具一方面可以及时发现社会中的深度伪造技术异化现象，以便对其进行针对性治理；另一方面还可以在社会中形成良好的监督响应机制，以对个体和组织的行为进行全方位监管。在国家安全层面，使用最多的政策工具是机构管制与监督机制，二者共同作用使国家的整体安全不受侵害。机构管制主要是对跨国机构及他国在本国的机构进行管控，预防其利用深度伪造技术散布不利于国家安全的虚假信息；而监督机制主要是对其他国家的深度伪造技术发展情况进行监督，防止本国在国际技术竞赛上落后于他国。在文化观念层面，运用较多的政策工具是宣传教育与信息安全文化，二者本质都是通过教育手段来加强人们的技术道德素养，潜移默化地影响人们的动机，实现“润物细无声”的技术异化治理。

5 国外深度伪造政策分析结论与启示

5.1 美国与欧盟深度伪造政策分析结论

通过分析 21 份深度伪造政策文本内容，笔者发现美国、欧盟的深度伪造政策在内容具有一定的特色和侧重点，但总体上都是在其现有制度的基础上增加对深度伪造技术异化现象的治理，以加强新政策与旧政策的相互促进和融合。

（1）深度伪造政策以环境型政策为主。由表 3 可知，深度伪造政策文本中环境型政策工具共有 177 个参考点，占全部编码参考点的 75%，由此可知，深度伪造政策以环境型政策为主。环境型政策工具对于限制和规范深度伪造技术的使用具有显著作用，例如，美国国会颁布的《深度伪造责任法案》是全球第一部针对深度伪造犯罪所制定的法案，其对深度伪造合成技术进行针对性打击，以遏制深度伪造技术使用者的恶意行为。再如，美国德克萨斯州政府制定的《关于制作欺骗性视频意图影响选举结果的刑事犯罪法案》与弗吉尼亚州政府颁布的《非法传播或出售他人影像》，这两部均属于明确且具有强制性的法案，其中规定涉嫌制作或传播深度伪造内容的人将罚款 5 000 美元或监禁两年半。综上可知，在当前复杂多变的互联网环境中，为对深度伪造技术异化进行有效治理，政府不仅需要制定引导性的环境型政策，还需要设计强制且硬性的环境型政策。

（2）深度伪造政策工具较为关注对技术、个体和组织的治理。从表 4 中可看出，美国与欧盟的深度伪造政策工具较多运用在技术、个体和组织上，主要原因在于以下方面：其一，技术异化大多是新技术所造成的不良影响，因而最为科学的治理方式是弥补技术的缺陷或创造出可与其对抗的新技术；其二，个体作为技术的使用者，其使用意图直接决定了技术的恶与善，因此对个体行为进行限制和管制是最为有效的治理方式；其三，组织机构在开发、研制新技术的过程中，具有技术伦理审查和监督的义务和责任，因此政府需对组织进行严格的管理，从开发者层面降低技术异化概率。

（3）深度伪造政策关注技术伦理的宣传与教育。从 X&Y 维度的分析中，可以发现已有政策工具作用于文化观念层面，主要是通过宣传教育使公众了解并合理使用技术，最终实现对其技术素养的提升。对于正在快速发展的国家和社会而言，技术伦理的宣传与教育是其技术市场稳定发展的“助推器”，一方面，技术伦理可引导技术主体在技术的设计、生产、使用等过程中充分考虑技术活动的附随后果，从而做出合乎伦理的选择；另一方面，当技术主体充分遵循技术伦理时，不仅会促进技术的创新发展，还会对国家和社会的发展带来正向的牵引。因此，在对深度伪造技术异化进行治理时，政府应当考虑如何进行技术伦理的宣传与教育，这样才能从源头上减少技术异化现象的发生。

5.2 中国深度伪造异化治理政策启示

在前文的政策结论基础之上，结合深度伪造技术在中国的应用以及技术异化治理的必要性，笔者得出以下政策启示。

（1）制定深度伪造技术异化治理方案。在复杂的网络环境中，为应对突发的深度伪造违法现象，政府应当制定技术异化治理方案：其一，针对深度伪造内容的恶意制作和传播行为制定相关法规管制及惩罚政策，目前中国的深度伪造恶意行为一般体现为未经他人同意的 AI 换脸、语音合成等行为，在一定程度上是对个人隐私权的侵犯，因此相关机构应当给予重视；其二，建立健全社会监督和反馈机制，运用监督机制政策工具加强社会中的深度伪造技术异化监督与反馈；其三，制定深度伪造技术标准体系，国家立法机关应与国标委合作制定深度伪造技术标准规范，对技术的开发与使用进行严格界定，以促进共同利益的实现。

（2）促进政府主导的深度伪造技术创新实践。由政府主导的技术创新实践不仅可以降低技术伦理带来的负面影响，还可以推动技术的创新创造。政府可利用技术扶持、鼓励创新、人才培养、科研活动等政策工具的组合实现上述目标。其中，科研活动政策强调政府与高校 / 科研机构的合作，主要形式是由政府出资、高校 / 科研机构负责技术研发。除此之外，政府可鼓励企业进行技术创新，并给予一定的政策支持，如技术扶持与税收优惠政策。

（3）实现技术伦理与政策工具的结合。当今人们生活在极为复杂的信息环境中，新兴技术带来了潜在风险、不确定以及多元价值选择等问题，因而技术伦理与政策的结合愈发变得迫切。技术伦理与政策工具结合指政府将技术发展的伦理考量纳入政策制定过程，通过一系列规约手段和规约机制更好地在实践中贯彻技术伦理要求，从而使技术发展更加符合伦理价值、满足社会需求[23]。为实现技术价值的最大化，政策应当嵌入深度伪造技术的设计、试验、应用、推广阶段，实现以政府为主导、多元主体共同参与的技术伦理治理机制，以及前瞻性的制度和规范建构。此外，政府可通过政策工具作用于舆论和教育，比如进行技术伦理的宣传教育、培养信息安全文化素养，将深度伪造技术伦理与社会价值相融合。

（4）积极参与国际深度伪造技术异化治理。世界局势暗流涌动，技术竞争一直都是国家之间的核心竞争所在。因此，中国不仅要关注国内技术异化的治理，还要积极参与国际中的技术异化治理，具体措施包括以下方面：其一，调查并评估其他国家深度伪造技术进展，比如美国深度伪造政策中就提出“对中国和俄罗斯使用深度伪造技术的情况进行评估”，由此看出，在维护国家安全同时也必须未雨绸缪，以保持本国利益的最大化；其二，参与国际协同治理机制建构，国际协同治理是一种国际组织通过“中间人”，利用软性约束开展间接治理的方式[24]，这种方式有助于实现全球技术异化治理。虽然国际协同治理由于种种原因很难实现，但中国作为世界大国之一应当主动开展多层次、多方位的国际合作，加深与国际组织的互动，适当调整原则并表达自身诉求，建立国际协同治理机制。

6 结　语

近年来，人工智能技术的飞速发展对社会造成了巨大而深远的影响，有人认为应该充分发挥技术的潜能来造福社会；但也有人极力抵触新兴技术的广泛使用，认为应该加以监管和限制。两种立场均言之有理，但为了避免技术异化对个人、社会和国家带来的潜在威胁，确保技术向善，政府应当提出强有力的举措来治理技术异化。本文聚焦于深度伪造技术的政策分析，通过收集、梳理、分析美国与欧盟的深度伪造政策文本内容，对政策文本中的政策工具及其所涉及的技术异化治理维度进行量化分析，从而结合中国国情提出关于深度伪造技术异化治理的政策建议。研究中还存在不足，如政策文本分析的样本是以美国和欧盟为例，因此在样本选择上存在一定的局限性，在进一步研究中，可扩大样本的选取范围，以增加政策分析的可信度。

参考文献：

[1] JUDGE H B, DIXON JR (Ret). Deepfakes: More frightening than

photoshop on steroids[J]. Judges' journal, 2019, 58(3): 35.

[2] GüERA D, DELP E J. Deepfake video detection using recurrent neural networks[C]. IEEE international conference on advanced video and signal based surveillance (AVSS), 2018(15): 1-6.

[3] 朱新同, 唐云祁, 耿鹏志. 基于特征融合的篡改与深度伪造图像检测算法[J]. 信息网络安全, 2021, 21(8): 70-81.
ZHU X T, TANG Y Q, GENG P Z. Falsification and deep forgery image detection algorithm based on feature fusion[J]. Information network security, 2021, 21(8): 70-81.

[4] 王丽娜, 聂建思, 汪润, 等. 面向深度伪造的溯源取证方法[J]. 清华大学学报(自然科学版), 2022, 62(5): 959-964.
WANG L N, NIE J S, WANG R, et al. A method of forensic tracing for deep forgery[J]. Journal of Tsinghua university (natural science edition), 2022, 62(5): 959-964.

[5] 张爱军, 王芳. 人工智能视域下的深度伪造与政治舆论变异[J]. 河海大学学报(哲学社会科学版), 2021, 23(4): 29-36, 106.
ZHANG A J, WANG F. Deepfake and the variation of political public opinion from the perspective of artificial intelligence[J]. Journal of Hohai university (philosophy and social sciences edition), 2021, 23(4): 29-36, 106.

[6] 龙坤, 马钺, 朱启超. 深度伪造对国家安全的挑战及应对[J]. 信息安全与通信保密, 2019(10): 21-34.
LONG K, MA Y, ZHU Q C. The challenge and response of deepfake to national security[J]. Information security and communication confidentiality, 2019(10): 21-34.

[7] CHESNEY R, CITRON D K. Deepfakes: A looming challenge for privacy, democracy, and national security[J]. California law review, 2019, 107(6): 1753-1819.

[8] 熊波. "深度伪造"的扩张化刑事治理风险及其限度[J]. 安徽大学学报(哲学社会科学版), 2020, 44(6): 106-113.
XIONG B. Risks and limits of expanded criminal governance of "deepfake"[J]. Journal of Anhui university (philosophy and social sciences edition), 2020, 44(6): 106-113.

[9] 陈昌凤, 徐芳依. 智能时代的 "深度伪造" 信息及其治理方式[J]. 新闻与写作, 2020(4): 66-71.
CHEN C F, XU F. "Deepfake" information in the age of intelligence and its governance[J]. News and writing, 2020(4): 66-71.

[10] CHESNEY R, CITRON D K. Deepfakes and the new disinformation war: The coming age of post-truth geopolitics[J]. Foreign affairs, 2019, 98(1): 147-155.

[11] 安全内参. ASPI 报告: 武器化深度造假对国家安全与民主的影响[EB/OL]. [2022-02-04]. https://www.secrss.com/articles/19188.
Security insider. ASPI report: The impact of weaponized deepfakes on national security and democracy[EB/OL]. [2022-02-04]. https://www.secrss.com/articles/19188.

[12] 安全内参. CSET 报告: Deepfakes 的基本威胁评估[EB/OL]. [2022-02-04]. https://www.secrss.com/articles/25428.
Security insider. CSET report: Basic threat assessment of Deepfakes [EB/OL]. [2022-02-04]. https://www.secrss.com/articles/25428.

[13] AMY M. Many Americans say made-up news is a critical problem that needs to be fixed[EB/OL]. [2022-02-04]. https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-need s-to-be-fixed/.

[14] 顾建光. 公共政策工具研究的意义、基础与层面[J]. 公共管理学报, 2006, 3(4): 58-61.
GU J G. The significance, basis and levels of research on public policy tools[J]. Journal of public administration, 2006, 3(4): 58-61.

[15] HOWLETT M, RAMESH M, PERL A. Studying public policy: Policy cycles and policy subsystems[M]. Oxford: Oxford university press, 2003: 92.

[16] PHHAL R, O'SULLIVAV E. A framework for mapping industrial emergence[J]. Technological forecasting and social change, 2011, 78(2): 217-230.

[17] ROTHWELL R, ZEGVELD W. Reindusdalization and technology[M]. Logman group limited, 1985: 83-104.

[18] 韩莉莉, 马万利. 技术异化视域下科技伦理人文效应探析[J]. 人民论坛·学术前沿, 2020(6): 92-95.
HAN L L, MA W L. An analysis of the humanistic effect of science and technology ethics from the perspective of technological alienation[J]. People's forum·academic frontiers, 2020(6): 92-95.

[19] 苗争鸣, 尹西明, 许展玮, 等. 颠覆性技术异化及其治理研究——以"深度伪造"技术的典型化事实为例[J]. 科学学与科学技术管理, 2020, 41(12): 83-98.
MIAO Z M, YIN X M, XU Z W, et al. Research on disruptive

technology alienation and its governance: Taking the typical fact of "deepfake" technology as an example[J]. Science and science and technology management, 2020, 41(12): 83–98.

[20] 贾婕. 生态社会主义视域下的科学技术异化[J]. 科学技术哲学研究, 2018, 35(3): 122–126.

JIA J. Alienation of science and technology from the perspective of ecological socialism[J]. Research in philosophy of science and technology, 2018, 35(3): 122–126.

[21] GRIMES M G, WILLIAMS T A, ZHAO E Y. Beyond hybridity: Accounting for the values complexity of all organizations in the study of mission and mission drift[J]. Academy of management review, 45 (1): 234–238.

[22] 柳卸林, 董彩婷, 丁雪辰. 数字创新时代: 中国的机遇与挑战[J]. 科学学与科学技术管理, 2020, 41(6): 3–15.

LIU X L, DONG C T, DING X C. The era of digital innovation: China's opportunities and challenges[J]. Science and science and technology management, 2020, 41(6): 3–15.

[23] 薛桂波, 闫坤如. “负责任创新”视角下技术伦理的政策转向[J]. 大连理工大学学报(社会科学版), 2018, 39(1): 9–14.

XUE G B, YAN K R. Policy turn of technological ethics from the perspective of "responsible innovation"[J]. Journal of Dalian university of technology (social science edition), 2018, 39(1): 9–14.

[24] 汤蓓. 试析国际组织的协同治理策略——以国际劳工组织推广“社会保障底线”政策为例[J]. 国际观察, 2017(3): 65–80.

TANG B. An analysis of the collaborative governance strategies of international organizations – Taking the promotion of the "social security bottom line" policy of the international labour organization as an example[J]. International observation, 2017(3): 65–80.

Analysis and Enlightenment of International Deepfake Technology Policy Texts: Taking the United States and the European Union as Examples

ZHAO Xueqin, LI Tian'e, HU Huihui

(School of History and Culture, Hubei University, Wuhan 430062)

Abstract: [Purpose/Significance] In the digital age, artificial intelligence (AI) technologies based on intelligent algorithms have a huge impact on people's lives, and one representative technology is deepfake technology. It is a derivative technology that appears naturally in the development of AI technology, and is mainly used in speech synthesis, AI face changing, etc. However, the abuse of deepfake technologies has caused varying degrees of threat to national security, social and public security, and privacy and reputation. Therefore, it is necessary for the government to formulate policies to guide the behavior of deepfake technology subjects. [Method/Process] Through online research, it is found that the United States and the European Union have paid much attention to the alienation of deepfake technology, and have formulated a series of policies to govern it. Therefore, the deepfake technology policies of the United States and the European Union were selected as the research object. This paper adopts the "X-Y" two-dimensional analysis framework to analyze the deepfake policy text. The X dimension is the basic policy tool, and the Y dimension is the perspective of how to deal with technology alienation. Finally, a comprehensive analysis of the two was carried out, that is, the application of policy tools in dealing with technology alienation is studied. This paper uses Nvivo12 to complete the coding of policy tools, codes the policy terms one by one according to the selected policy tool types, and then analyzes the coding results and draws conclusions. [Results/Conclusions] The

deepfake technology policies of the United States and the European Union have certain characteristics in content, such as focusing on environmental policy tools, paying more attention to the governance of technology, individuals and organizations, and focusing on the publicity and education of technology ethics. But in general, it is based on its existing system to increase the governance of the alienation phenomenon of deepfake technology, so as to strengthen the mutual promotion and integration of new policies and old policies. For China, the governance of deepfake technology alienation should be started as soon as possible. Specific measures include formulating deepfake technology alienation governance plans, promoteing government-led deepfake technology innovation practices, realizing the combination of technology ethics and policy tools, actively participating in international governance of deepfake technology alienation. In addition, this paper takes the deepfake technology policies of the United States and the European Union as an example, which has certain limitations, so in the future research, we will expand the selection of samples.

Keywords: deepfake; policy analysis; technological alienation; national security; disinformation